



#5

residues and the carat indicates the mismatched base site (G). Figure 1(C) provides a schematic plasmid indicating the sequence of the kan chimeric double-stranded hairpin oligonucleotide (left; SEQ ID NO: 4341) and the sequence the tet chimeric double-stranded hairpin oligonucleotide used in other experiments (SEQ ID NO: 4342). Figure 1(D) provides a flow chart of a kan experiment in which a chimeric double-stranded hairpin oligonucleotide (SEQ ID NO: 4341) is used. In Figure 1(D), the Kan mutant sequence corresponds to SEQ ID NO: 4343 and SEQ ID NO: 4344; the Kan converted sequence corresponds to SEQ ID NO: 4345 and SEQ ID NO: 4346; the mutant sequence in the sequence trace corresponds to SEQ ID NO: 4347 and the converted sequences in the sequence trace correspond to SEQ ID NO: 4348.

Figure 2. *Genetic readout system for correction of a point mutation in plasmid pK^sm4021.* A mutant kanamycin gene harbored in plasmid pK^sm4021 is the target for correction by oligonucleotides. The mutant G is converted to a C by the action of the oligo. Corrected plasmids confer resistance to kanamycin in *E.coli* (DH10B) after electroporation leading to the genetic readout and colony counts. The wild type sequence corresponds to SEQ ID NO: 4349.

Figure 3: *Target plasmid and sequence correction of a frameshift mutation by chimeric and single-stranded oligonucleotides.* (A) Plasmid pT^sΔ208 contains a single base deletion mutation at position 208 rendering it unable to confer tet resistance. The target sequence presented below indicates the insertion of a T directed by the oligonucleotides to re-establish the resistant phenotype. (B) DNA sequence confirming base insertion directed by Tet 3S/25G; the yellow highlight indicates the position of frameshift repair. The wild type sequence corresponds to SEQ ID NO: 4350, the mutant sequence corresponds to SEQ ID NO: 4351 and the converted sequence corresponds to SEQ ID NO: 4352. The control sequence in the sequence trace corresponds to SEQ ID NO: 4353 and the 3S/25A sequence in the sequence trace corresponds to SEQ ID NO: 4354.

Figure 4. *DNA sequences of representative kan^r colonies.* Confirmation of sequence alteration directed by the indicated molecule is presented along with a table outlining codon distribution. Note that 10S/25G and 12S/25G elicit both mixed and unfaithful gene repair. The number of clones sequenced is listed in parentheses next to the designation for the single-stranded oligonucleotide. A plus (+) symbol indicates the codon identified while a figure after the (+) symbol indicates the number of colonies with a particular sequence. TAC/TAG indicates a mixed peak. Representative DNA sequences are presented below the table with yellow highlighting altered residues. The sequences in the sequence traces have been assigned numbers as follows: 3S/25G, 6S/25G and 8S/25G correspond to SEQ ID NO: 4355, 10S/25G corresponds to SEQ ID NO: 4356, 25S/25G on the lower left corresponds to SEQ ID NO: 4357 and 25S/25G on the lower right corresponds to SEQ ID NO: 4358.

Figure 5. *Gene correction in HeLa cells.* Representative oligonucleotides of the invention are co-transfected with the pCMVneo(+)FIAsH plasmid (shown in Figure 9) into HeLa cells. Ligand is diffused into cells after co-transfection of plasmid and oligonucleotides. Green fluorescence indicates gene correction of the mutation in the antibiotic resistance gene. Correction of the mutation results in the expression of a fusion protein that carries a marker ligand binding site and when the fusion protein binds the ligand, a green fluorescence is emitted. The ligand is produced by Aurora Biosciences and can readily diffuse into cells enabling a measurement of corrected protein function; the protein must bind the ligand directly to induce fluorescence. Hence cells bearing the corrected plasmid gene appear green while "uncorrected" cells remain colorless.

5

Figure 6. *Z-series imaging of corrected cells*. Serial cross-sections of the HeLa cell represented in Figure 5 are produced by Zeiss 510 LSM confocal microscope revealing that the fusion protein is contained within the cell.

Figure 7. *Hygromycin-eGFP target plasmids*. (A) Plasmid pAURHYG(ins)GFP contains a single base insertion mutation between nucleotides 136 and 137, at codon 46, of the Hygromycin B coding sequence (cds) which is transcribed from the constitutive ADH1 promoter. The target sequence presented below indicates the deletion of an A and the substitution of a C for a T directed by the oligonucleotides to re-establish the resistant phenotype. In Figure 7A, the sequence of the normal allele corresponds to SEQ ID NO: 4359, the sequence of the target/existing mutation corresponds to SEQ ID NO: 4360 and the sequence of the desired alteration corresponds to SEQ ID NO: 4361. (B) Plasmid pAURHYG(rep)GFP contains a base substitution mutation introducing a G at nucleotide 137, at codon 46, of the Hygromycin B coding sequence (cds). The target sequence presented below the diagram indicates the amino acid conservative replacement of G with C, restoring gene function. In Figure 7B, the sequence of the normal allele corresponds to SEQ ID NO: 4359, the sequence of the target/existing mutation corresponds to SEQ ID NO: 4362 and the sequence of the desired alteration corresponds to SEQ ID NO: 4361.

Figure 8. *Oligonucleotides for correction of hygromycin resistance gene*. The sequence of the oligonucleotides used in experiments to assay correction of a hygromycin resistance gene are shown. DNA residues are shown in capital letters, RNA residues are shown in lowercase and nucleotides with a phosphorothioate backbone are capitalized and underlined. In Figure 8, the sequence of HygE3T/25 corresponds to SEQ ID NO: 4363, the sequence of HygE3T/74 corresponds to SEQ ID NO: 4364, the sequence of HygE3T/74a corresponds to SEQ ID NO: 4365, the sequence of HygGG/Rev corresponds to SEQ ID NO: 4366 and the sequence of Kan70T corresponds to SEQ ID NO: 4367.

Figure 9. *pAURNeo(-)FIAsH plasmid*. This figure describes the plasmid structure, target sequence, oligonucleotides, and the basis for detection of the gene alteration event by fluorescence. In Figure 9, the sequence of the Neo/kan target mutant corresponds to SEQ ID NO: 4343 and SEQ ID NO: 4344, the converted sequence corresponds to SEQ ID NO: 4345 and SEQ ID NO: 4346 and the FIAsH peptide sequence corresponds to SEQ ID NO: 4368.

Figure 10. *pYESHyg(x)eGFP plasmid*. This plasmid is a construct similar to the pAURHyg(x)eGFP construct shown in Figure 7, except the promoter is the inducible GAL1 promoter. This promoter is inducible with galactose, leaky in the presence of raffinose, and repressed in the presence of dextrose.

The following examples are provided by way of illustration only, and are not intended to limit the scope of the invention disclosed herein.

EXAMPLE 1
Assay Method For Base Alteration
And Preferred Oligonucleotide Selection

In this example, single-stranded and double-hairpin oligonucleotides with chimeric backbones (see Figure 1 for structures (A and B) and sequences (C and D) of assay oligonucleotides) are used to correct a point mutation in the kanamycin gene of pK^sm4021 (Figure 2) or the tetracycline gene of pT^sΔ208 (Figure 3). All kan oligonucleotides share the same 25 base sequence surrounding the target base identified for change, just as all tet oligonucleotides do. The sequence is given in Figures 1C and Figure 1D. Each plasmid contains a functional ampicillin gene. Kanamycin gene function is restored

when a G at position 4021 is converted to a C (via a substitution mutation); tetracycline gene function is restored when a deletion at position 208 is replaced by a C (via frameshift mutation). A separate plasmid, pAURNeo(-)FIAsH (Figure 9), bearing the *kan^s* gene is used in the cell culture experiments. This plasmid was constructed by inserting a synthetic expression cassette containing a neomycin phosphotransferase (kanamycin resistance) gene and an extended reading frame that encodes a receptor for the FIAsH ligand into the pAUR123 shuttle vector (Panvera Corp., Madison, WI). The resulting construct replicates in *S. cerevisiae* at low copy number, confers resistance to aureobasidinA and constitutively expresses either the Neo+/FIAsH fusion product (after alteration) or the truncated Neo-/FIAsH product (before alteration) from the ADH1 promoter. By extending the reading frame of this gene to code for a unique peptide sequence capable of binding a small ligand to form a fluorescent complex, restoration of expression by correction of the stop codon can be detected in real time using confocal microscopy. Additional constructs can be made to test additional gene alteration events.

We also construct three mammalian expression vectors, pHyg(rep)eGFP, pHyg(Δ)eGFP, pHyg(ins)eGFP, that contain a substitution mutation at nucleotide 137 of the hygromycin-B coding sequence. (rep) indicates a T137 \Rightarrow G replacement, (Δ) represents a deletion of the G137 and (ins) represents an A insertion between nucleotides 136 and 137. All point mutations create a nonsense termination codon at residue 46. We use pHygEGFP plasmid (Invitrogen, CA) DNA as a template to introduce the mutations into the hygromycin-eGFP fusion gene by a two step site-directed mutagenesis PCR protocol. First, we generate overlapping 5' and a 3' amplicons surrounding the mutation site by PCR for each of the point mutation sites. A 215 bp 5' amplicon for the (rep), (Δ) or (ins) was generated by polymerization from oligonucleotide primer HygEGFPf (5'-AATACGACTCACTATAGG-3'; SEQ ID NO: 4369) to primer Hygrepr (5'-GACCTATCCACGCCCTCC-3'; SEQ ID NO: 4370), Hyg Δ r (5'-GACTATCCACGCCCTCC-3'; SEQ ID NO: 4371), or Hyginsr (5'-GACATTATCCACGCCCTCC-3'; SEQ ID NO: 4372), respectively. We generate a 300bp 3' amplicon for the (rep), (Δ) or (ins) by polymerization from oligonucleotide primers Hygrefp (5'-CTGGGATAGGTCCTGCGG-3'; SEQ ID NO: 4373), Hyg Δ f (5'-CGTGGATAGTCCTGCGG-3'; SEQ ID NO: 4374), Hyginsf (5'-CGTGGATAATGTCCTGCGG-3'; SEQ ID NO: 4375), respectively to primer HygEGFPf (5'-AAATCACGCCATGTAGTG-3'; SEQ ID NO: 4376). We mix 20 ng of each of the resultant 5' and 3' overlapping amplicon mutation sets and use the mixture as a template to amplify a 523 bp fragment of the Hygromycin gene spanning the KpnI and RsrII restriction endonuclease sites. We use the Expand PCR system (Roche) to generate all amplicons with 25 cycles of denaturing at 94°C for 10 seconds, annealing at 55°C for 20 seconds and elongation at 68°C for 1 minute. We digest 10 μ g of vector pHygEGFP and 5 μ g of the resulting fragments for each mutation with KpnI and

- 17A -

RsrII (NEB) and gel purify the fragment for enzymatic ligation. We ligate each mutated insert into pHygEGFP vector at 3:1 molar ratio using T4

contained plasmid molecules harboring the targeted base correction. While a few colonies appeared on plates derived from reaction mixtures containing 25-mers with 10 or 12 thioate linkages on both ends, the sequences of the plasmid molecules from these colonies contain nonspecific base changes. In these illustrative examples, the second base of the codon is changed (see Figure 3). These results show that modified single-strands can direct gene repair, but that efficiency and specificity are reduced when the 25-mers contain 10 or more phosphorothioate linkages at each end.

In Figure 1, the numbers 3, 6, 8, 10, 12 and 12.5 respectively indicate how many phosphorothioate linkages (S) or 2'-O-methyl RNA nucleotides (R) are at each end of the exemplified molecule although other molecules with 2, 4, 5, 7, 9 and 11 modifications at each end can also be tested. Hence oligo 12S/25G represents a 25-mer oligonucleotide which contains 12 phosphorothioate linkages on each side of the central G target mismatch base producing a fully phosphorothioate linked backbone, displayed as a dotted line. The dots are merely representative of a linkage in the figure and do not depict the actual number of linkages of the oligonucleotide. Smooth lines indicate DNA residues, wavy lines indicate 2'-O-methyl RNA residues and the carat indicates the mismatched base site (G).

Correction of a mutant kanamycin gene in cultured mammalian cells. The experiments are performed using different mammalian cells, including, for example, 293 cells (transformed human primary kidney cells), HeLa cells (human cervical carcinoma), and H1299 (human epithelial carcinoma, non-small cell lung cancer). HeLa cells are grown at 37°C and 5% CO₂ in a humidified incubator to a density of 2 x 10⁵ cells/ml in an 8 chamber slide (Lab-Tek). After replacing the regular DMEM with Optimem, the cells are co-transfected with 10 µg of plasmid pAURNeo(-)FIAsH and 5 µg of modified single-stranded oligonucleotide (3S/25G) that is previously complexed with 10 µg lipofectamine, according to the manufacturer's directions (Life Technologies). The cells are treated with the liposome-DNA-oligo mix for 6 hrs at 37°C. Treated cells are washed with PBS and fresh DMEM is added. After a 16-18 hr recovery period, the culture is assayed for gene repair. The same oligonucleotide used in the cell-free extract experiments is used to target transfected plasmid bearing the kan^s gene. Correction of the point mutation in this gene eliminates a stop codon and restores full expression. This expression can be detected by adding a small non-fluorescent ligand that bound to a C-C-R-E-C-C (SEQ ID NO: 4385) sequence in the genetically modified carboxy terminus of the kan protein, to produce a highly fluorescent complex (FIAsH system, Aurora Biosciences Corporation). Following a 60 min incubation at room temperature with the ligand (FIAsH-EDT2), cells expressing full length kan product acquire an intense green fluorescence detectable by fluorescence microscopy using a fluorescein filter set. Similar experiments are performed using the HygeGFP target as described in Example 2 with a variety of mammalian cells, including, for

subject to inhibition by RNA residues than by phosphorothioate linkages. Thus, even though both of these oligonucleotides contain an equal number of modifications to impart nuclease resistance, XI (with 16 phosphorothioate linkages) has good gene repair activity while VII (with 16 2'-O-methyl RNA residues) is inactive. Hence, the original chimeric double hairpin oligonucleotide enabled correction directed, in large part, by the strand containing a large region of contiguous DNA residues.

Oligonucleotides can target multiple nucleotide alterations within the same template. The ability of individual single-stranded oligonucleotides to correct multiple mutations in a single target template is tested using the plasmid pK^sm4021 and the following single-stranded oligonucleotides modified with 3 phosphorothioate linkages at each end (indicated as underlined nucleotides): Oligo1 is a 25-mer with the sequence TTCGATAAGCCTAT**GCTGACCCG**TG (SEQ ID NO: 4377) corrects the original mutation present in the kanamycin resistance gene of pK^sm4021 as well as directing another alteration 2 basepairs away in the target sequence (both indicated in boldface); Oligo2 is a 70-mer with the 5'-end sequence TTCGGCTACGACTGGGCACAACAGACAATTGGC (SEQ ID NO: 4378) with the remaining nucleotides being completely complementary to the kanamycin resistance gene and also ending in 3 phosphorothioate linkages at the 3' end. Oligo2 directs correction of the mutation in pK^sm4021 as well as directing another alteration 21 basepairs away in the target sequence (both indicated in boldface).

We also use additional oligonucleotides to assay the ability of individual oligonucleotides to correct multiple mutations in the pK^sM4021 plasmid. These include, for example, a second 25-mer that alters two nucleotides that are three nucleotides apart with the sequence 5'-TTGTGCCCAGTC**G**TATCCG AATAGC-3' (SEQ ID NO: 4379); a 70-mer that alters two nucleotides that are 21 nucleotides apart with the sequence 5'-CATCAGAGCAGCC**A**ATTGTCTGTTGTGCCCAGTC**G**TAGCCGAATAGCCTCTCCACCCAA GCGGCCGGAGA-3' (SEQ ID NO: 4380); and another 70-mer that alters two nucleotides that are 21 nucleotides apart with the sequence 5'-GCTGACAGCCGGAACACGGCGGCATCAGAGCAGCC**A**ATTG TCTGTTGTGCCCAGTC**G**TAGCCGAATAGCCT-3' (SEQ ID NO: 4381). The nucleotides in the oligonucleotides that direct alteration of the target sequence are underlined and in boldface. These oligonucleotides are modified in the same way as the other oligonucleotides of the invention.

We assay correction of the original mutation in pK^sm4021 by monitoring kanamycin resistance (the second alterations which are directed by Oligo2 and Oligo3 are silent with respect to the kanamycin resistance phenotype). In addition, in experiments with Oligo2, we also monitor cleavage of the resulting plasmids using the restriction enzyme Tsp509I which cuts at a specific site present only when the second alteration has occurred (at ATT in Oligo2). We then sequence these clones to

Example 1 indicating that oligonucleotides comprising phosphorothioate linkages facilitate gene correction much more efficiently than control duplex, chimeric RNA-DNA oligonucleotides. This gene correction activity is also specific as transformation of cells with the control oligonucleotide Kan70T produced no hygromycin resistant colonies above background and thus Kan70T did not support gene correction in this system. In addition, we observe that the 74-base oligonucleotide (HygE3T/74) corrects the mutation in pAURHYG(ins)GFP approximately five-fold more efficiently than the 25-base oligonucleotide (HygE3T/25). We also perform control experiments with LSY678 yeast cells containing the plasmid pAURHYG(wt)GFP. With this strain we observed that even without added oligonucleotides, there are too many hygromycin resistant colonies to count.

We also use additional oligonucleotides to assay the ability of individual oligonucleotides to correct multiple mutations in the pAURHYG(x)eGFP plasmid. These include, for example, one that alters two basepairs that are 3 nucleotides apart is a 74-mer with the sequence 5'-CTCGTGCTTTCA GCTTCGATGTAGGAGGGCGTGGG**TAC**GTCTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTAC-3' (SEQ ID NO: 4382); a 74-mer that alters two basepairs that are 15 nucleotides apart with the sequence 5'-CTCGTGCTTT**CAG**CTTCGATGTAGGAGGGCGTGGATA**CGT**CCTGCGGGTAA**AC**AGCTGCGCCGATG GTTTCTAC-3' (SEQ ID NO: 4383); and a 74-mer that alters two basepairs that are 27 nucleotides apart with the sequence 5'-CTCGTGCTTT**CAG**CTTCGATGTAGGAGGGCGTGGATA**CGT**CCTGCGGGTAAAT AGCTGCGCCG**AC**GGTTTCTAC (SEQ ID NO: 4384). The nucleotides in these oligonucleotides that direct alteration of the target sequence are underlined and in boldface. These oligonucleotides are modified in the same ways as the other oligonucleotides of the invention.

Oligonucleotides targeting the sense strand direct gene correction more efficiently. We compare the ability of single-stranded oligonucleotides to target each of the two strands of the target sequence of both pAURHYG(ins)GFP and pAURHYG(rep)GFP. These experiments, presented in Tables 7 and 8, indicate that an oligonucleotide, HygE3T/74 α , with sequence complementary to the sense strand (i.e. the strand of the target sequence that is identical to the mRNA) of the target sequence facilitates gene correction approximately ten-fold more efficiently than an oligonucleotide, HygE3T/74, with sequence complementary to the non-transcribed strand which serves as the template for the synthesis of RNA. As indicated in Table 7, this effect was observed over a range of oligonucleotide concentrations from 0-3.6 μ g, although we did observe some variability in the difference between the two oligonucleotides (indicated in Table 7 as a fold difference between HygE3T/74 α and HygE3T/74). Furthermore, as shown in Table 8, we observe increased efficiency of correction by HygE3T/74 α relative to HygE3T/74 regardless of whether the oligonucleotides were used to correct the base substitution

Clinical Phenotype & Mutation	Correcting Oligos	SEQ ID NO:
	GAATCAAAAGCAAACTTCCTCTGACTCTATCTTCATTATTTGG GAGCTTATCATTGAAGTCCTTGGAATTATTTTCAAATTCTGTT TCTTTGAATCTTTGTTGTCTGAGAAACTCTCT	1762
	TTCCAAGGACTTCAATG	1763
	CATTGAAGTCCTTGAA	1764
Adenomatous polyposis coli Leu2839Phe CTT-TTT	AAAACTGACAGCACAGAATCCAGTGGAACCCAAAGTCCTAAG CGCCATTCTGGGTCTTACCTTGAGATCTGTTTAAAGAGAG GAAGAATGAACTAAGAAAATTCTATGTTAATTACA	1765
	TGTAATTAACATAGAATTTCTTAGTTTCATTCTTCCTCTCTTT AAACAGATGTCACAAGGTAAGACCCAGAATGGCGCTTAGGAC TTTGGGTTCCACTGGATTCTGTGCTGTCAGTTTT	1766
	GGTCTTACCTTGAGACA	1767
	TGTCACAAGGTAAGACC	1768

EXAMPLE 12
Parahemophilia - Factor V Deficiency

Deficiency in clotting Factor V is associated with a lifelong predisposition to thrombosis. The disease typically manifests itself with usually mild bleeding, although bleeding times and clotting times are consistently prolonged. Individuals that are heterozygous for a mutation in Factor V have lowered levels of factor V but probably never have abnormal bleeding. A large number of alleles with a range of presenting symptoms have been identified. The attached table discloses the correcting oligonucleotide base sequences for the Factor V oligonucleotides of the invention.

Table 19
Factor V Mutations and Genome-Correcting Oligos

Clinical Phenotype & Mutation	Correcting Oligos	SEQ ID NO:
Factor V deficiency Ala221Val GCC-GTC	TTGACTGAATGCTTATTTGGCCTGTGTCTCTCCCTCTTTCTCA GATATAACAGTTTGTGCCCATGACCACATCAGCTGGCATCTGC TGGGAATGAGCTCGGGGCCAGAATTATTCTCCAT	4340
	ATGGAGAATAATTCTGGCCCCGAGCTCATTCCCAGCAGATGC CAGCTGATGTGGTCATGGGCACAACTGTTATATCTGAGAAAG AGGGAGAGACACAGGCCAAAATAAGCATTCAAGTCAA	1769
	AGTTTGTGCCCATGACC	1770
	GGTCATGGGCACAACT	1771
Thrombosis Arg306Gly AGG-GGG	TGTCCTAACTCAGCTGGGATGCAGGCTTACATTGACATTAAAA ACTGCCCAAAGAAAACCAAGGAATCTTAAGAAAATAACTCGTGA GCAGAGGCGGCACATGAAGAGGTGGGAATACTTCA	1772